



NONLINEAR TRAVELTIME OPTIMIZATION FOR RAY TRACING IN COMPLEX 3D MEDIA

Danilo R. Velis

Departamento de Geofísica Aplicada, Facultad de Cs. Astronómicas y Geofísicas,
Universidad Nacional de La Plata,
Paseo del Bosque s/n, La Plata 1900, Argentina
(Tel +54 21 21-7308, E-mail velis@fcaglp.edu.ar)

ABSTRACT

A new method for solving the boundary value ray tracing problem in laterally variable media is presented. The strategy overcomes some well known difficulties that arise in standard shooting and bending methods. Problems related to: (1) the selection of new take-off angles, and (2) local minima in multipathing cases, are overcome by using a stochastic optimization algorithm. At each iteration, the ray is propagated from the source by solving a standard initial value problem. The last portion of the raypath is then forced to pass through the receiver. The problem is put into a nonlinear optimization scheme in which total traveltimes is globally minimized with respect to the initial conditions that produce the absolute minimum path. A number of examples of multipathing in 2D and 3D media are examined.

INTRODUCTION

Ray tracing plays an important role in seismological studies. Large attention has been devoted to the *initial value problem (IVP)*, in which the ray is specified by the initial conditions: initial point and initial direction of propagation (*take-off angles*). The IVP is in general a well resolved problem [1]. However, geotomographic methods and earthquake location usually require precise traveltimes and trajectory computations of seismic waves propagating between two fixed points in an laterally heterogeneous medium. This represents a *boundary value problem (BVP)* because the ray is not only specified by the initial conditions.

Traditionally, there are two methods for solving the two-point BVP: shooting and bending [2] [3]. The shooting method represents a standard IVP. First, an initial point (*source*) is fixed and the ray is propagated by specifying the take-off angles. Then, a search strategy is used to update these angles until the ray emerges through the desired endpoint (*receiver*). Since frequently the receiver location is an ill-behaved function of the take-off angles, the strategy for choosing the new take-off angles may become a difficult task, and divergence is as common issue. The problem is even more severe in the 3D case, where two take-off angles are to be found.

In the bending method both points are linked by an initial guess path, which is then perturbed iteratively so as to satisfy the ray equations. Several bending techniques are reported in the literature [2, 4, 5, 6], which, unlike shooting, always produce a ray connecting any source-receiver

pair. In general, bending involves the solution of a highly nonlinear optimization problem, which requires some kind of gradient directions to update the raypath. In complicated velocity structures, bending tends to overlook multipath propagation because the solution depends on the first guess. As a result, the absolute minimum raypath can be missed.

I present here a new method for ray tracing through general 3D media that I call Simulated Annealing Ray Tracing (SART). The purpose is to overcome the usual deficiencies of shooting and bending for solving the BVP. The philosophy of the SART method is to put the BVP into a convenient optimization framework which is in turn solved by means of an efficient simulated annealing algorithm. In the two-point ray tracing case, SART is an iterative procedure that attempts to find the optimum take-off angles corresponding to the raypath with minimum traveltimes connecting any given source-receiver pair [7].

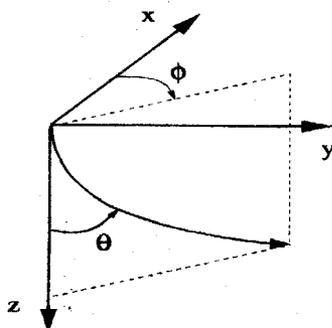


Fig. 1. The ray direction is described by declination θ and azimuth ϕ .

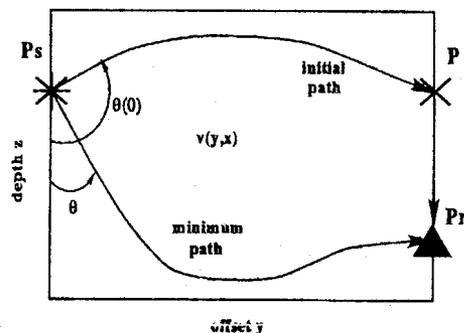


Fig. 2. Straight line construction used by SART. When traveltimes is minimum, P coincides with P_r and the minimum path with take-off angle θ has been found.

THE INITIAL VALUE PROBLEM (IVP)

The ray equations are well developed in the literature [8, 1]. In their final form, they may be written

$$\begin{cases} \partial_t x = v \sin \theta \cos \phi, \\ \partial_t y = v \sin \theta \sin \phi, \\ \partial_t z = v \cos \theta, \\ \partial_t \theta = -\cos \theta \left(\frac{\partial v}{\partial x} \cos \phi + \frac{\partial v}{\partial y} \sin \phi \right) + \frac{\partial v}{\partial z} \sin \theta, \\ \partial_t \phi = \frac{1}{\sin \theta} \left(\frac{\partial v}{\partial x} \sin \phi - \frac{\partial v}{\partial y} \cos \phi \right). \end{cases} \quad (1)$$

where θ and ϕ stand for the azimuth and declination angles which describe the direction of the ray at every point of its trajectory (Figure 1), and $v = v(x, y, z)$ is the wavespeed. Given the initial conditions, the system (1) can be solved numerically (4th-order Runge-Kutta is a good choice) and thus it is possible to describe the ray trajectory at every time t , the independent variable of integration. The initial conditions ($t = 0$) are usually given by

$$\begin{cases} [x(0), y(0), z(0)] = (x_s, y_s, z_s) \\ \theta(0) = \theta_0 \\ \phi(0) = \phi_0, \end{cases} \quad (2)$$

where subscript s stands for *source*.

MODEL REPRESENTATION

The numerical integration requires the right-hand side of (1) to be continuous and v to be twice differentiable. To handle discontinuities, I describe the model by a number of regions separated by surface interfaces. Rays are propagated until a discontinuity is found where the reflection/transmission law is applied.

The velocity model is compounded of any number of regions separated by surface interfaces representing geologic horizons or other discontinuities (e.g. fractures). I assume all surface interfaces are one-to-one functions given by $z = g(x, y)$. The velocity within each region may be specified by any function $v = v(x, y, z)$, and must be twice differentiable. This model representation is quite flexible to describe a wide variety of velocity structures.

THE BOUNDARY VALUE PROBLEM (BVP)

A two-point BVP must be solved when both the initial (source) and the ending (receiver) points are specified. Although the solution sometimes is nonunique, the purpose of the following method is to find the raypath whose traveltimes is a global minimum.

Simulated Annealing Ray Tracing (SART)

Basically, at each iteration an IVP is solved starting from the source, P_s , with take-off angles (θ_0, ϕ_0) . The propagation is terminated provided either: (1) the ray arrives at the model boundary; or (2) the ray arrives at a prescribed near-receiver region, e.g. some predefined target plane.

The point where the ray meets one of the above two conditions, is called P . The raypath is completed by connecting P with the receiver, P_r , with a straight line. Figure 2 illustrates this situation when condition (1) is met for a 2D model. Note the segment that connects P with P_r is quite arbitrary, and the resulting raypath may be completely unrealistic. But this is an intermediate raypath that, like in bending, is updated iteratively.

The total traveltimes is computed by integrating the path length s between both endpoints:

$$T = \int_{P_s}^{P_r} \frac{1}{v} ds = \int_{P_s}^P \frac{1}{v} ds + \int_P^{P_r} \frac{1}{v} ds. \quad (3)$$

Since source and receiver are fixed and P is uniquely determined by the solution of the IVP [I assume the IVP can be solved for any given pair of take-off angles (θ_0, ϕ_0)], T becomes a function of the take-off angles only, so

$$T = T(\theta_0, \phi_0). \quad (4)$$

The final raypath is found by recalling Fermat's principle. When traveltimes T is minimum, P coincides with P_r and the whole raypath satisfies the ray equations. (As a matter of fact this is not always strictly true, and some ray for which $P \neq P_r$ may arrive earlier than any other *realistic* paths. However, I will reformulate the problem so as to overcome this eventual difficulty).

The BVP now becomes an optimization problem in which two parameters (take-off angles) are to be found so that T is a global minimum. Since expression (4) is a highly nonlinear, multimodal and nondifferentiable function, it cannot be properly minimized using classical linearizing methods. I use instead a global optimization algorithm called Very Fast Simulated Annealing (VFSA) [9], which is a very powerful tool for minimizing arbitrary functions independently of its nonlinearities, discontinuities and stochasticity.

Boundary value ray tracing as a constrained optimization problem

The boundary value ray tracing problem can be viewed as a constrained global optimization problem. In the two-point ray tracing case, the function to be globally minimized is the traveltime from P_s to P , and the constraint is that P must coincide with P_r , within a given tolerance. The set of parameters that satisfy the constraint is called *feasible region*. Put it mathematically,

$$\text{minimize } T_1 = \int_{P_s}^P \frac{1}{v} ds, \quad \text{subject to } d \leq d_{tol}, \quad (5)$$

where d is the distance between P and P_r and d_{tol} is a tolerance distance. Since usually the traveltime represents an observation with a certain error T_{tol} , the optimization problem can be rewritten

$$\text{minimize } T_1 = \int_{P_s}^P \frac{1}{v} ds, \quad \text{subject to } T_2 = \int_P^{P_r} \frac{1}{v} ds \leq T_{tol}. \quad (6)$$

Not only problem (6) is nonlinear (except for some simple form of the velocity field, for which the path is known) but also multimodal and nondifferentiable in general. This is a rather difficult optimization problem.

In standard shooting methods, the optimization problem is put simply as

$$\text{minimize } d = \|P - P_r\|. \quad (7)$$

Although this formulation seems much simpler, it is necessary to find all minima of d for which $d \leq d_{tol}$ to decide later which one corresponds to the minimum traveltime. Note some minima may not belong to the feasible region (local minima).

In SART I solve (6) using VFSA, taking special care to properly incorporate the constraint. For this purpose, I transform the constrained optimization problem (6) into an unconstrained one by defining the following cost function:

$$T(\theta_0, \phi_0) = T_1 + \rho T_2, \quad (8)$$

where $\rho \geq 1$ is a penalty factor. In practice, it is generally enough to set $\rho = 1$ for the global minimizer to be a feasible solution.

OPTIMIZATION USING SIMULATED ANNEALING (SA)

In 1953, Metropolis et al presented a Monte Carlo sampling technique for modeling the evolution of a solid at a given temperature [10]. In 1983, Kirkpatrick et al generalized the concept and applied it to nonlinear optimization problems [11]. Here the unknown parameters play the role of the *particles* in the solid, and the cost function represents the *energy* of the system. In the SA approach, at each iteration the parameter space is randomly perturbed and new configurations are accepted (or rejected) so that the cost function decreases. Occasionally, some increases of the cost function are allowed so that the system can escape from local minima.

The decision for accepting/rejecting new configurations is known as the *Metropolis criterion* [10]. If ΔE_k is the change of the cost function at iteration k with respect to the previous iteration, the criterion states that a new configuration is accepted unconditionally if $\Delta E_k < 0$, and accepted with probability

$$h = \exp(-\Delta E_k/c_k) \quad (9)$$

if $\Delta E \geq 0$, where c_k is the *temperature* of the system at the current iteration.

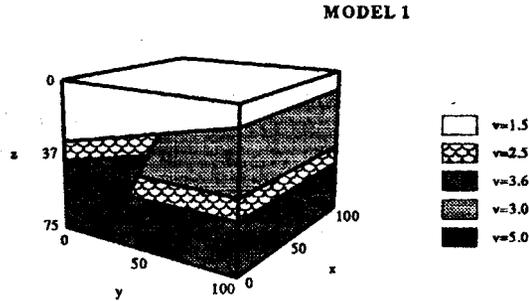


Fig. 3. Blocky model with planar interfaces and constant velocities. A fracture plane has also been defined. Coordinates are given in units and velocities in units per second.

Table 1. Multipath in MODEL 1

Take off angle	Traveltime [s]
70.16	35.986
87.44	36.410
108.21	35.915
125.90	34.714

Initially, the temperature is high enough so that almost all proposed configurations are accepted. At low temperatures, the probability of accepting a new configuration corresponding to an increase of the cost function is small. Clearly, as the temperature approaches zero, the acceptance probability decreases exponentially and only the lowest energy states are accepted. The system is stopped after a maximum number of iterations or when no further improvement in the cost function is observed.

The selection of an appropriate cooling schedule is essential in SA. In [12] it is showed that a global minimum of E can be obtained (statistically) provided the temperature is lowered no faster than $c_k = c_0 / \ln(k)$, where c_0 is a constant. For many practical applications, this cooling schedule is too slow, and many researchers use a faster cooling schedule, but global convergence is no longer guaranteed. These types of algorithms are called *simulated quenching* (SQ) rather than SA [13]. However, a faster cooling schedule can indeed be used that still guarantees convergence by modifying the generating function that governs the parameter perturbation stage. In [9] it is proposed a VFSA technique that allows a much faster cooling schedule.

In VFSA parameters are drawn from a long-tailed Cauchy-like distribution that expands and contracts dynamically according to the sensitivity of the cost function of each dimension (this procedure is called "re-annealing"). At high temperatures, the model space is sampled more or less uniformly, but at low temperatures, models with lowest energy are preferably sampled. The long-tailed distribution permits exploration of the parameter space more effectively, and a faster cooling rate is allowed to accelerate convergence. In VFSA, the temperature is given by

$$c_{i,k} = c_{i,0} \exp(-\alpha_i k^{1/M}) \quad i = 1, M \quad (10)$$

where M is the number of parameters and α_i is some *user-defined* constant.

The advantage of VFSA over traditional SA techniques lies in the choice of the generating distribution and the cooling schedule, which has proven that VFSA is faster than other SA algorithms [15].

TEST RESULTS

Two blocky models representing geological structures were considered. Distances and coordinates are expressed in distance units, velocities in units per second and traveltimes in seconds. Angles are in degrees. Model 1 is compounded of 7 regions with constant velocities delimited by plane interfaces and a fracture plane, as shown in Figure 3. For simplicity, I first take into

account a 2D slice of the model: the plane $x = 50$. I placed a source at $(50, 0, 70)$ and a receiver at $(50, 100, 16.5)$, and produced a fan of rays with take-off angles θ varying from 60 to 140 degrees (ϕ is fixed at 90 degrees so that all ray trajectories lie on the plane $x = 50$). Figure 4 shows distance d and traveltimes T as a function of θ , with $\rho = 1$. The nonlinearity of the objective functions and the complexity of the optimization problem are evident by observing the plots. Local minima, multipathing and discontinuities are all present in this simple two-dimensional example. Distance d , for example, has four zeros (multipathing) and two extra local minima, among various discontinuities generated by the velocity structure. Clearly, the multipathing problem cannot be obviated by minimizing d . The shape of curve T is similar to the shape of curve d , but the former makes it possible to differentiate among those rays arriving in the receiver. These raypaths are shown in Figure 5. Table 1 summarizes their corresponding take-off angle and traveltimes.

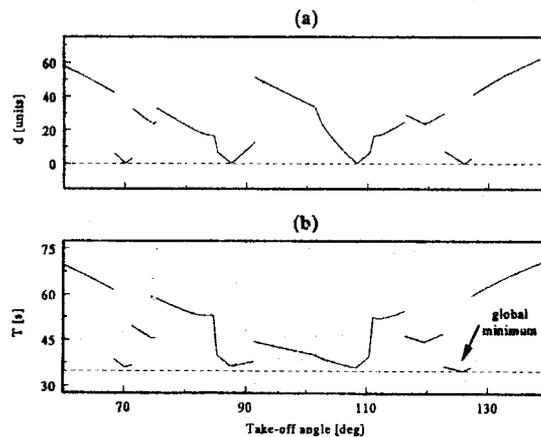


Table 2. Multipath in MODEL 2

Take off angle	Traveltime [s]
41.15	43.567
41.84	43.566
43.89	41.805
44.75	43.588
50.03	41.718
53.79	39.717
58.61	46.985
59.50	46.652
78.02	49.780
80.07	49.748

Fig. 4. (a) Distance d as a function of take-off angle θ .
(b) Traveltime T as a function of take-off angle θ . Both curves correspond to Model 1.

In the previous 2D example, an exhaustive scan of all possible take-off angles is viable in order to find the absolute minimum raypath. But in a 3D model, where raypaths are specified by more than one parameter, a scan is not recommended for obvious reasons. Figure 6 shows T as a function of θ and ϕ for Model 1 (3D version). What makes it difficult to globally minimize T is not only the presence of local minima, but also the great number of discontinuities generated by a blocky model. Simulated annealing appears to be a natural tool for solving this kind of nonlinear optimization problems. In effect, convergence was achieved after 200 iteration as shown in Figure 7. Note that both θ and ϕ were involved in the optimization.

Model 2 represents a salt dome with several layers with non-constant velocities (Figure 8). A total of eight regions are delimited by nonplanar interfaces with cylindrical symmetry. The same symmetry is not observed for all the velocities. A source is located at $(0, -50, 0)$ and a receiver at $(0, 50, 0)$. Again, for simplicity consider $\phi = 90$ so that all raypaths will lie on the plane $x = 0$. Figure 9 shows d and T as a function of θ . The complexity of these two functions is enormous: nine rays arrive in the receiver satisfying the ray equations, more than 10 other local minima and an uncountable number of discontinuities are also present in function d . Figure 10 shows these nine raypaths and Table 2 summarizes their traveltimes and take-off angles.

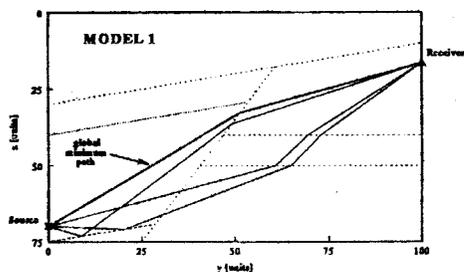


Fig. 5. Multipathing in Model 1. The four plotted raypaths satisfy the ray equations, but exhibit different traveltimes.

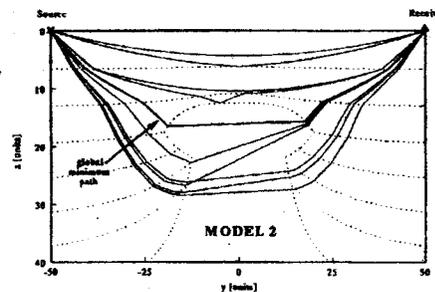


Fig. 10. Multipathing in Model 2. The nine plotted raypaths satisfy the ray equations, but exhibit different traveltimes.

By inspecting the curves in Figure 9, it is clear that any local minimizer would likely fail to converge to a useful solution. Function d cannot be used because it is impossible to differentiate among the nine raypaths that connect source and receiver. On the contrary, the global minimum of T corresponds to the desired raypath. Besides, the global minimum lies in a very narrow valley of less than one degree. If ϕ is also considered, the topography of $T(\theta, \phi)$ becomes so complex I was not able to plot it at all. Despite the rather difficult optimization problem, SART found the true solution in about 300 iterations, as shown in Figure 7.

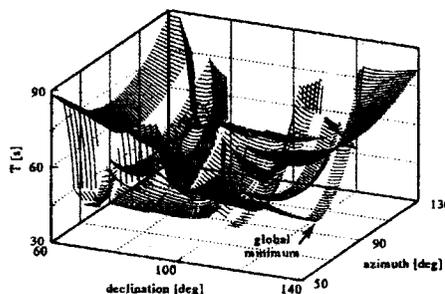


Fig. 6. Traveltime T as a function of both take-off angles. Note the complex topography. The surface is explored during the SA process in order to locate the global minimum.

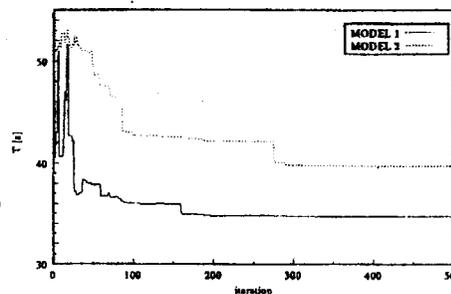


Fig. 7. Convergence behavior of SART for Model 1 and Model 2. Each iteration involves basically the solution of one IVP.

CONCLUSIONS

I have presented a new method for solving the two-point ray tracing problem which exhibits some improvements over existing techniques. Specifically, the problem of local minimum path which exists in the bending method is fully overcome, as well as the difficulties associated with the strategy for choosing the appropriate take-off angles in the shooting method. It is emphasized that, unlike in bending methods, the physical nature of the resulting ray is known a priori and a posteriori, a point that is very important in phase identification. In SART, any available initial value ray tracer can be used to generate the raypath at each iteration. Moreover, the accuracy of the results is not dependent on the parameterization of the ray trajectory, but on the selected initial value ray tracer algorithm. One possible disadvantage is that raypaths

connecting each source-receiver pair are obtained one at a time. However, since the dimension of the search space is very low, the computational cost is not at all related to the accuracy of the ray parameterization, like in bending methods.

Even for moderately complex 3D structures, the associated nonlinear optimization problem which is involved in solving the BVP represents a very difficult task for standard techniques. SART makes use of a powerful stochastic optimization algorithm to find the solution which exhibits the global minimum traveltimes. The results are encouraging because good results were obtained despite the fact that traveltimes may be an extremely ill-behaved function of the take-off angles.

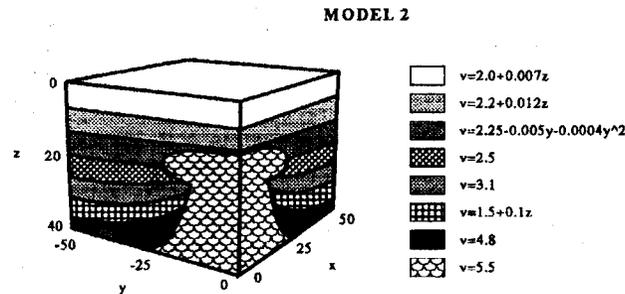


Fig. 8. Blocky model with curved interfaces and non constant velocities. Note the salt dome with high velocity contrast. Coordinates are given in units and velocities in units per second.

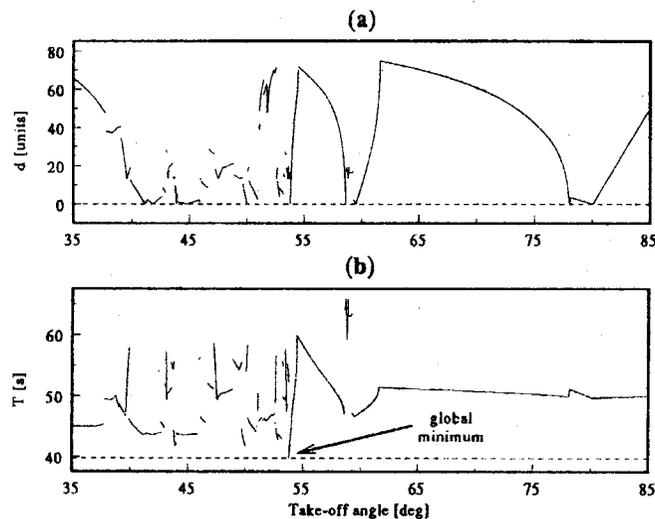


Fig. 9. (a) Distance d as a function of take-off angle θ . (b) Traveltime T as a function of take-off angle θ . Both curves correspond to Model 2. Note how ill-behaved the curves are. A great number of local minima and discontinuities make them extremely difficult to minimize using standard methods. Besides the global minimum of T lies in a very narrow valley.

REFERENCES

- [1] Červený, V., 1987: Ray tracing algorithms in three-dimensional laterally varying layered structures, in *Seismic Tomography*, 99-133, ed. Nolet, G., Reidel, Dordrecht.
- [2] Julian, B.R., and Gubbins, D., 1977, Three-dimensional seismic ray tracing, *J. Geophys.*, *43*, 95-113.
- [3] Aki, K. and Richards, P.G., 1980: *Quantitative seismology: theory and methods*, W.H. Freeman and Co.
- [4] Um, J., and Thurber, C., 1987, A fast algorithm for two-point seismic ray tracing, *Bull., Seis. Soc. Am.*, *77*, 972-986.
- [5] Prothero, W.A., Taylor, W.J., and Eickemeyer, J.A., 1988, A fast, two-point, three dimensional ray tracing algorithm using a simple step search method, *Bull., Seis. Soc. Am.*, *78*, 1190-1198.
- [6] Moser, T.J., Nolet, G., and Snieder, R., 1992, Ray bending revisited, *Bull., Seis. Soc. Am.*, *82*, 259-288.
- [7] Velis, D.R. and Ulrych, T.J., 1996, Simulated annealing two-point ray tracing, *Geophysical Research Letters*, *23*, 201-204.
- [8] Eliseevnin, V.A., 1965, Analysis of rays propagating in an inhomogeneous medium, *Sov. Phys. Acoust. (English Translation)*, *10*, 242-245.
- [9] Ingber, L., 1989, Very fast simulated re-annealing: *Mathl. Comput. Modelling*, *12*, 967-973.
- [10] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., 1953, Equation of state calculations by fast computing machines, *J. Chem. Phys.* *21*, 1087-1092.
- [11] Kirkpatrick, S., Gellat, C.D. Jr. and Vecchi, M.P., 1983, Optimization by simulated annealing, *Science*, *220*, 671-680.
- [12] Geman, S. and Geman, D., 1984, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images: *IEEE Trans. Pattern Analysis and Machine Intelligence*, *PAMI-6*, 721-741.
- [13] Ingber, L., 1993, Simulated annealing: practice versus theory, *Mathl. Comput. Modelling*, *18*, 11, 29-57.
- [14] Rosen, B., 1992, Function optimization based on advanced simulated annealing: Presented at IEEE Workshop on Physics and Computation, Expanded Abstracts, 289-293.

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in the context of public administration and government operations. This section outlines the various methods and systems used to collect, store, and analyze data, ensuring that information is readily accessible and reliable.

2. The second part of the document focuses on the challenges and solutions associated with data management. It identifies common issues such as data redundancy, inconsistency, and security concerns. The text provides detailed recommendations for addressing these challenges, including the implementation of robust data governance policies, the use of advanced data management tools, and the establishment of clear roles and responsibilities for data stewards.

3. The third part of the document explores the integration of data across different departments and systems. It highlights the benefits of a unified data environment, such as improved decision-making, enhanced operational efficiency, and better service delivery. The text discusses various integration strategies, including data warehousing, data lakes, and API-based architectures, and provides examples of successful implementation cases.

4. The fourth part of the document addresses the ethical and legal implications of data collection and analysis. It discusses the importance of obtaining informed consent, ensuring data privacy, and complying with relevant regulations such as the General Data Protection Regulation (GDPR). The text also touches upon the potential for bias and discrimination in data-driven decision-making and provides guidance on how to mitigate these risks.

5. The fifth part of the document discusses the future of data management and analytics. It explores emerging technologies such as artificial intelligence, machine learning, and big data, and their potential to revolutionize data analysis and decision-making. The text also discusses the importance of ongoing education and training for data professionals to stay current in a rapidly evolving field.

6. The sixth part of the document provides a summary of the key findings and recommendations. It reiterates the importance of a data-driven approach and the need for a comprehensive data management strategy. The text concludes with a call to action, encouraging organizations to embrace data as a strategic asset and to implement the recommended practices to maximize their data's potential.